

The diagram illustrates the primary structure of insulin, a single-chain polypeptide consisting of 51 amino acids. The sequence is arranged in a circular fashion, starting from the N-terminus (labeled 1) and ending at the C-terminus (labeled 51). The amino acids are connected by peptide bonds, forming a continuous chain. Two interchain disulfide bonds (S-S) are shown, connecting the two chains of the insulin molecule. The amino acids are labeled as follows:

- 1: Gly
- 2: Leu
- 3: Gly
- 4: Lys
- 5: Ala
- 6: Gln
- 7: Cys
- 8: Ala
- 9: Leu
- 10: Trp
- 11: Leu
- 12: Ala
- 13: Cys
- 14: Gln
- 15: Ala
- 16: Dha
- 17: Gly
- 18: Gly
- 19: Gly
- 20: Ile
- 21: Gly
- 22: Ala
- 23: Gly
- 24: Gly
- 25: Val
- 26: Ala
- 27: Cys
- 28: Gln
- 29: Asn
- 30: Tyr
- 31: Arg
- 32: Gln
- 33: Phe
- 34: Cys
- 35: Arg
- 36: COO<sup>-</sup>
- 37: +H<sub>3</sub>N

Figure 1

# Sublancin Display Peptide

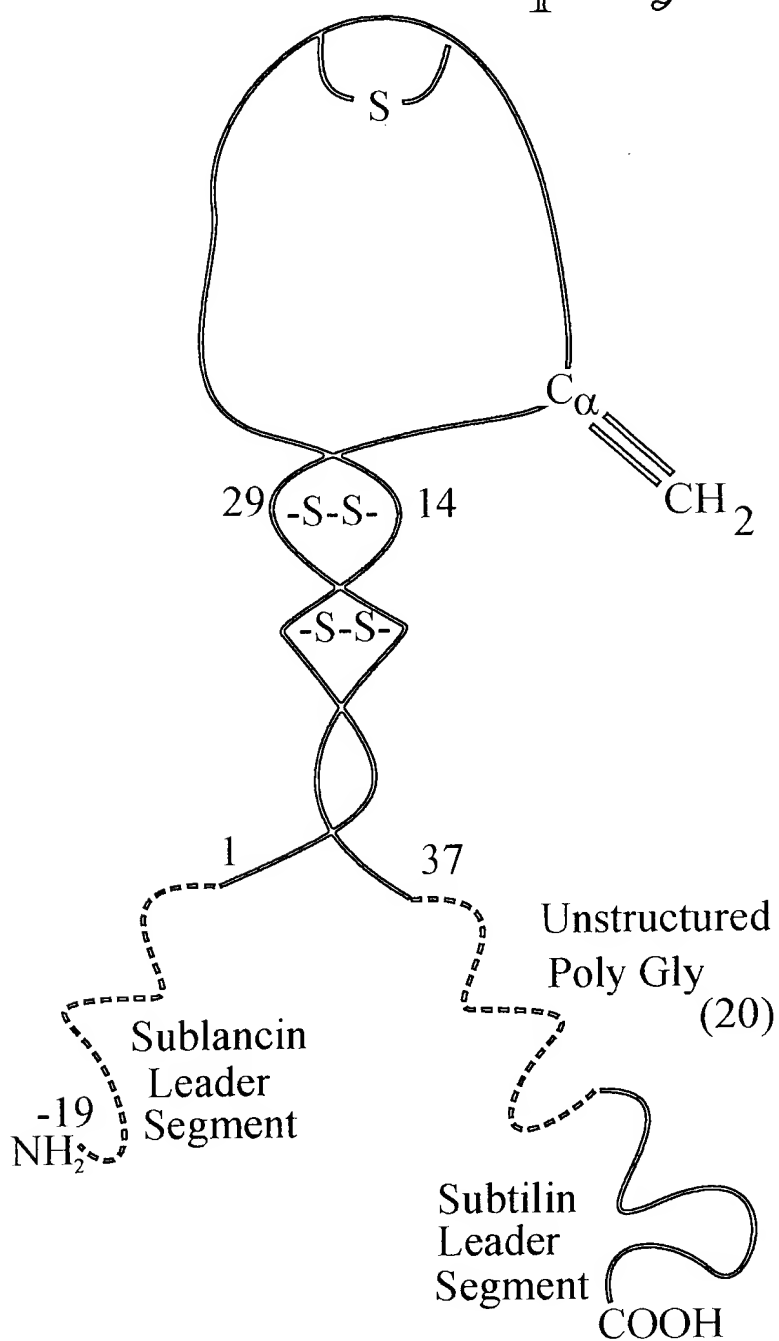


Figure 2

AGAAGTGTCTCAGTCACGTTATCGAATATTGAGGATGATGTTAATCAGCAGCTGAGTTTATTTGAAGTGG  
 A'TAATGAAAAAGAGAAGGAAACTCGGTTTTGTAAATGGATGGGATTAGAAGTAAATACGGCTCTAAAGCGAT  
 LPHF1--->  
 TCTGAGAGCAGTTTCTTATACACCAGCAGGAACTGCACTTCAACGAGCTGGATTAAACAGGTGGGCATAAG  
 AGTTAAGATAAATTTAACTTATATAACACATCGCTTAAAGTTTTTTGTTTTTAAAAACTTAAAAAACAT  
 |-----> *yolF* ----->  
 GGTAAAATTATATAAAAAACATAAGAAAGAGTGATTAT ATGGAATATGTAGTTATGATAATCATTTTATTA  
 GCACTTTTCTTTATTTTACTGTTTTCTTAAATACACGTTATAGTTTTGATGAAAAATGCTTAGTCTTAA  
 AATTTGGTTTATCTAAAAACAGAAATTCCAATTAATCAAATAGTTAGTATTAAAGAGTCAGACAAGTATGG  
 AGTTGCAGATAATATCGATTATAAAATTGGTATGCCATATGCTCAACCAGATAGAATTGTTATTGAAACT  
 <----- *yolF* <-----<br>
 ACAAATAAGCGTTTTCTAGTTTTTTTTTAAATGGAGCTCAACAATTTATTCAAAGTATAAAAGGGTTAGTG  
 --|  
 TT TGAACATAAAAAAGTACCTTCTTACAATAGAAGGTAAGTCTTTTTTGTATCTATAATTATTAAAAATTTAC  
 CTAAATTTTATCATTATTAATTCAAATAAATCCATAATAGTCAATTTTATTAGTGTATTACAACCAA  
 <---LPHR1, (LPHF2, LPVF2-->  
 TTCTGTTTATTGATAGGTAATAAAGTTTTTTTTCTATGATTTATGAACAAAGTTTCCTTATAATTTTCAA  
 -35 -10  
 AAAAAATAAAAAATATGGTTGAATTTAGATTTATCTTCCTTTATATTAAAAAATGTAATCCGGATTGCAA  
 r.b.s. |-----> *sunA* leader region----->  
 ACAAATGGGGAGGTTTTACAA ATGGAAGCTATTAAAGAAGTTAACTAGAGGAACTCGAAAAACCAA  
 <---LPHR2 <---LPVR2 NLPVF3----->  
*sunA* mature region ----->  
 AAGGTAGT GGATTAGGAAAAGCTCAGTGTGCTGCGTTGTGGCTACAATGTGCTAGTGGCGGTACAATTGG  
 <----- *sunA* <----- Pst I |  
 TTGTGGTGGCGGAGCTGTTGCTTGTCAAAAC TATCGTCAATTCTGCAGA TAAACATTTGTAGAGGGAAT  
 LPVF4---> LPHF3--->  
 <---LPPMR2  
 |-----> *sunT* ----->  
 ATTTTAAATATTCCTCATATTTAAAGCGGGGATTGAAA TTGAATAAGAAAAAGAAATATGTTCTACTA  
 AACAGTTTAATAGTCATGATTGTGGACTAGCTTGTATCTCGTCAATTTTAAAGTTTCATAACCTTAACTA  
 TGGAAATTGATTCTTACTAGACCTAATTGGGGATAAGGAAGGCTATAGTTTAAAGAGACTTAATTGTTATT  
 TTTAAGAAGATGGGGATAAAAACTAGGCCACTTGAATTGCAAGAAAAATAAGACATTGGAAGCCCTAAAC  
 AAATAAAGCTCCCTTGTATAGCTTTGTTAGAAGGGGAGGAATATGGACATTACATAACAATATACGAAAT  
 TAGAAATAACTATTTACTTGTAGTGATCCTGATAAAGACAAAATAACTAAAAATAAAAAAGAGGATTTT  
 GAAAGTAAATTCACAACTTTATATTAGAAATTGACAAAGAGTCAATTCCTGAAAAAGAAAAAGATCAAA  
 AAAACATTCTTACTTTTTTAAGGACATACTTTTAGAAATAAATTGATCGTTTTTGTGATTTTATTGAC  
 TTCCTTGTTCGTTGTGGGTCTTGTCTGCTAGCTGGGTCGTTTTATATAAAGTTTCTAGTTGACCT----->  
 <---LPHR3 & LPVR4 -----> *sunT* ----->

Figure 3

EcoRI  
 pTZ sequence <-----GAATTCGGCTCTAAAGCGAT  
 TCTGAGAGCAGTTTCTTATACACCAGCAGGAAGTGCCTTCAACGAGCTGGATTAAACAGGTGGGCATAAG  
 AGTTAAGATAAAATTTAACTTATATAACACATCGCTTAAAGTTTTTTTGTTTTTAAAACTTAAAAACAT  
 GGTAAAATTATATAAAAAACATAAGAAAAGAGTGATTATATGGAATATGTAGTTATGATAATCATTTTATTA  
 GCACTTTTCTTTATTTTTACTGTTTTTCCTAAATACACGTTATAGTTTTGTATGAAAAATGCTTAGTCTTAA  
 AATTTGGTTTATCTAAAACAGAAATTTCCAATTAATCAAATAGTTAGTATTAAAGAGTCAGACAAGTATGG  
 AGTTGCAGATAATATCGATTATAAAATTGGTATGCCATATGCTCAACCAGATAGAATTGTTATTGAACT  
 ACAAATAAGCGTTTTCTAGTTTTTTTTTAAATGGAGCTCAACAATTTATTCAAAGATATAAAAGGGTTAGTG  
 TTTGAACATAAAAAAGTACCTTCTTACAATAGAAGGTACTTTTTTGTATCTATAATTATTAAAAATTTAC  
 CTAATTTTTTATCATTATTAATTCAAAATAAATCCATAATAGTCAATTTTATTTAGTGTATTACAACCAA  
 Bam HI ( ~900 bp ) Bam HI  
 TTC GGATCC <----cat----> GGATTCGTGTATTACAACCAATTC TGTTTATTGATAGGTAATAAA  
 GTTTTTTTTCTATGATTTATGAACAAGTTTCCTTATAATTTTCAA  
 AAAAAATAAAAAATATGGTTGAATTTAGATTTATCTTCCTTTATATTAAAAAATGTAATCCGGATTGCAA  
 | Sublancin leader -----> Xho I  
 ACAAATGGGGAGGTTTTACAA ATGGAAAAGCTATTTAAAGAAGTTAACTCGAGGAACTCGAAAACCAA  
 | Sun A ----->  
 AAGGTAGT GGATTAGGAAAAGCTCAGTGTGCTGCGTTGTGGCTACAATGTGCTAGTGGCGGTACAATTGG  
 Pst I |  
 TTGTGGTGGCGGAGCTGTTGCTTGTCAAACTATCGTCAATTCTGCAGA TAAACATTTGTAGAGGGAAT  
 ATTTTAAATATTCCCTCATATTTAAAGCGGGGATTGAAATTGAATAAGAAAAAGAAATATGTTCATACTA  
 AACAGTTTAAATAGTCATGATTGTGGACTAGCTTGTATCTCGTCAATTTTAAAGTTTCATAACCTTAACTA  
 TGGAATTGATTTCTTACTAGACCTAATTGGGGATAAGGAAGGCTATAGTTTAAAGAGACTTAATTGTTATT  
 TTTAAGAAGATGGGGATAAAAACTAGGCCACTTGAATTGCAAGAAAATAAGACATTGGAAGCCCTAAAAAC  
 AAATAAAGCTCCCTTGTATAGCTTTGTTAGAAGGGGAGGAATATGGACATTACATAACAATATACGAAAT  
 TAGAAATAACTATTTACTTGTAGTGATCCTGATAAAGACAAAAATAACTAAAATAAAAAAGAGGATTTT  
 GAAAGTAAATTACAACTTTTATATTAGAAATTGACAAAGAGTCAATTCCTGAAAAAGAAAAAGATCAAA  
 AAAACATTCTTACTTTTTTAAAGGACATACTTTTTTAGAAATAAATTGATCGTTTTTGTGATTTTATTGAC  
 TTCCTTGTTCGTTGTGGGTCTTGCTGAAGCTT----->pTZ sequence  
 HindIII

### Figure 4

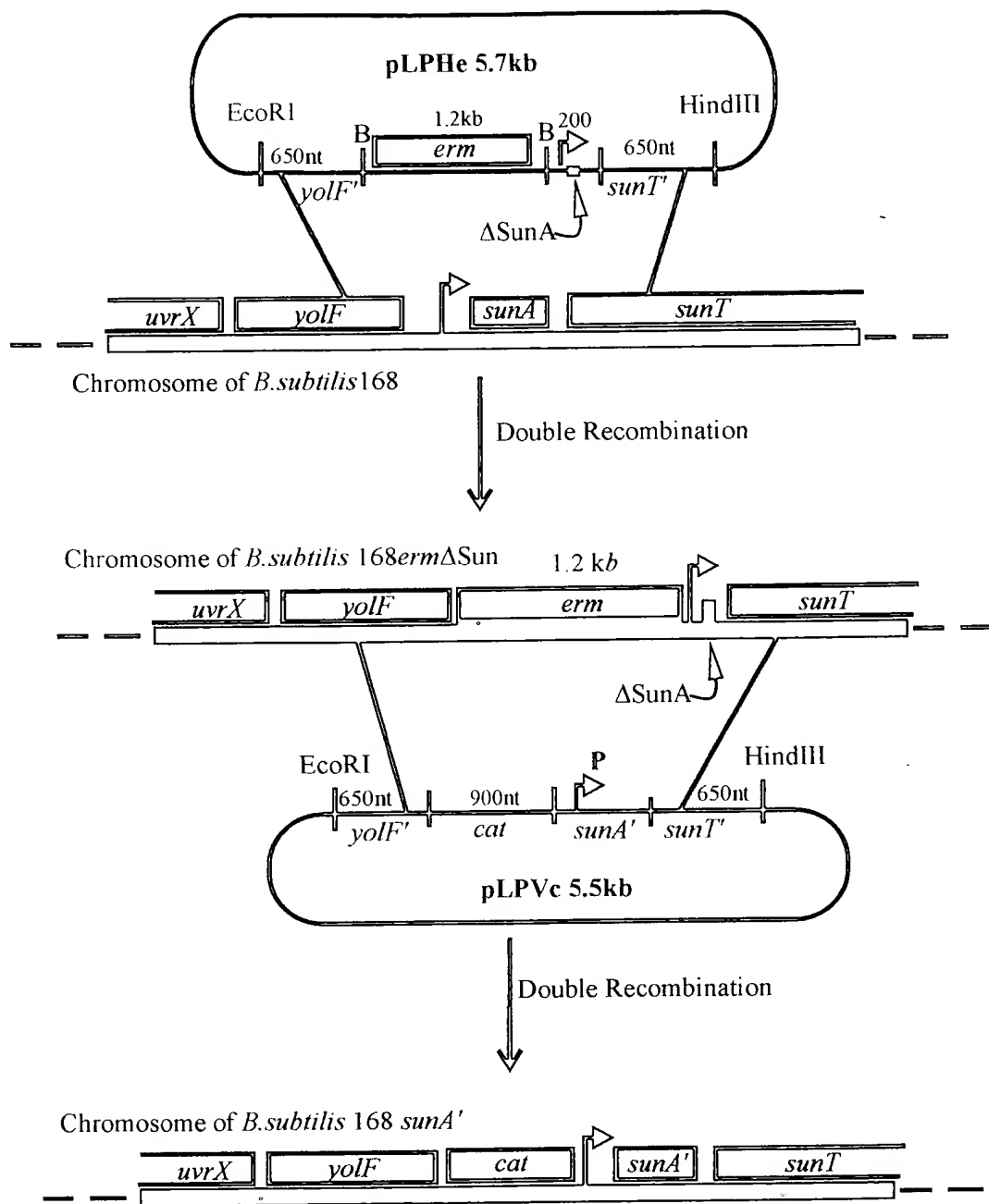
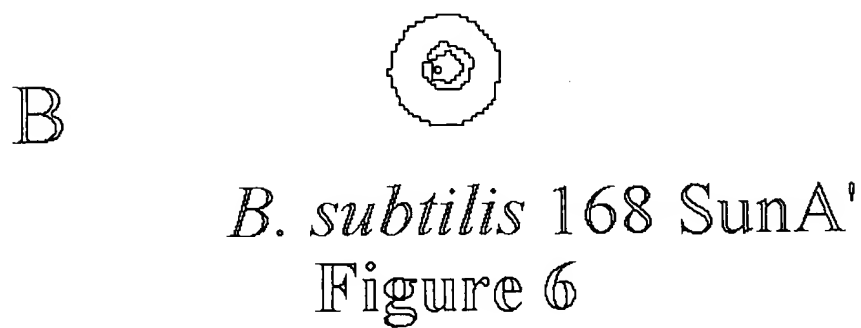
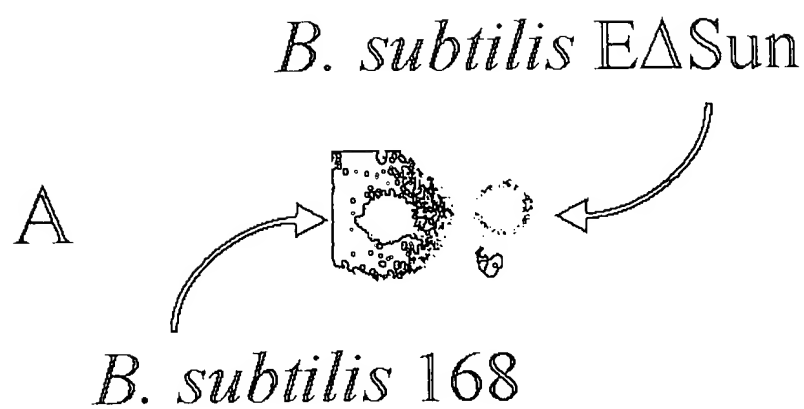


Figure 5



← pLPcat Sublancin leader→  
TTGCAAACAAATGGGGAGGTTTTACAA ATGGAAAAGCTATTTAAGAAG  
MetGluLysleuPheLysGluV

XhoI sublancin prep-  
TTAAACTCGAGGAAGTTCGAAAACCAAAAAGGTAGT GGATTAGGAAAAGC  
AllyLeuGluGluLeuGluAsnGluLysGlySer GlyLeuGlyLysAl  
tide→  
TCAGTGCTGCTGCGTTGTGGCTACAATGTGCTAGTGGCGGTACAATTGGTT  
aGlnCysAlaAlaLeuTrpLeuGlnCysAlaSerGlyGlyThrIleGlyC

KasI Poly-  
GTGGTGGCCGCCGCTTGTCTCAAACTATCGTCAATTCTGTAGAGGT  
ysGlyGlyGlyAlaValAlaCysGlnAsnTyrArgGlnPheCysArgGly

glycine20→ BseRI  
GGTGGTGGGGGAGGCCGGGGGAGGGGTGGTGGTGGAGGAGGTGGTGGTGG  
GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGl

subtilin leader→ XbaI  
TGGTGGTATGTCAAAGTTCGATGATTTTCGATCTAGATGTTGTGAAAGTCT  
yGlyGlyMetSerLysPheAspAspPheAspLeuAspValValLysVals

Stop PstI  
CTAAACAAGACTCAAAAATCACTCCGCAATAGAGTCCTGCAGATAAAACA  
erLysGlnAspSerLysIleThrProGln \* pLPcat →

Figure 7

